

Agreement on the evaluation of glaucomatous optic nerve head findings by ophthalmology residents and a glaucoma specialist

Julia D Rossetto
Luiz Alberto S Melo Jr
Mauro S Campos
Ivan M Tavares

Department of Ophthalmology and
Visual Sciences, Paulista School of
Medicine, Universidade Federal de São
Paulo, São Paulo, Brazil

Objectives: To assess agreement among ophthalmology residents and a glaucoma expert in the evaluation of cross-sectional glaucomatous optic nerve head characteristics using stereoscopic photographs.

Methods: Twenty stereo photographs were analyzed by ophthalmology residents just after completion of their first (First-Year Group) or third (Third-Year Group) year of residency and by a glaucoma expert. The agreement was assessed using the kappa statistic (κ) and limits of agreement.

Results: Agreement among resident groups and the expert ranged from poor to moderate. Agreement between Third Years and the expert seems to be better than that between First Years and the expert, especially in the evaluation of “nasal cupping”, “barring circumlinear vessel”, “notching”, and “retinal nerve fiber layer defect” criteria. However, no improvement was seen in the agreement with the expert regarding glaucomatous optic neuropathy, which was 64% ($\kappa=0.19$) for First Years and 63% ($\kappa=0.20$) for Third Years.

Conclusion: Agreement between residents and the expert was poor to moderate and similar when comparing both groups. This may suggest that the residents learn how to identify glaucoma signals during the first year of training, and the results of this study may facilitate the creation of targeted teaching tools in residency training.

Keywords: optic disk, residency and internship, medical education, glaucoma

Introduction

Glaucoma is an optic neuropathy that is associated with progressive loss of visual function. It is the leading cause of irreversible blindness worldwide.¹ One of the most important signs of progression of glaucomatous damage is the change in the appearance of the optic disk. Therefore, assessment of the optic disk is important for early detection, monitoring, and management of glaucoma. Stereoscopic color photography of the optic disk is used in the clinic to record the appearance of the optic nerve head (ONH). It is the gold standard method for detecting ONH and peripapillary retinal nerve fiber layer (RNFL) changes.

Other authors have reported high reproducibility and ability to detect optic disk changes.^{2,3} However, the intraobserver agreement is known to be better than interobserver agreement. Moreover, interobserver agreement improves with observer experience.⁴

In this study, we assessed interobserver agreement for the evaluation of ONH parameters on the stereo, color, fundus photographs between a glaucoma expert and ophthalmology residents after the first year (First-Year Group) and after the third year (Third-Year Group) of ophthalmology residency.

Correspondence: Ivan M Tavares
Department of Ophthalmology and
Visual Sciences, Paulista School of
Medicine, Universidade Federal de São
Paulo, Rua Botucatu, 821 São Paulo,
SP 04023-062, Brazil
Fax +55 11 5082 3599
Email glaucomaclinik@gmail.com

Materials and methods

The study was conducted in the Department of Ophthalmology and Visual Sciences of the Universidade Federal de São Paulo, São Paulo, Brazil. This study was approved by the Institutional Ethics Committee and complied with the principles outlined in the Declaration of Helsinki. After a full explanation of the procedures involved in the study, written informed consent was obtained from all participants.

Stereophotographic exams of the ONH were used in this prospective study. A random sample of high-quality, simultaneous stereoscopic photographs from a glaucoma clinic retinographer was selected by another glaucoma specialist, and these photos were from a database that included healthy patients and patients at all stages of glaucoma. No specific optic disk characteristic was chosen in a predetermined manner in this selection. Simultaneous color stereoscopic disk photographs were taken with a Visucam (Carl Zeiss Meditec, Dublin, CA, USA).

All ophthalmology residents ($n=13$) who had completed the first year (First-Year Group) and all ($n=13$) who had completed the third year (Third-Year Group) of residency training were simultaneously invited to participate in this study. No additional training on stereo photographs analysis was performed by the residents before completion of the assignment. Study participants assessed 20 stereo pairs of disk photographs on a computer monitor in a dark room with the help of stereo glasses (ScreenVu – Berezin Stereo Photography Products, Mission Vision, CA, USA) and were required to fill a form to determine the presence of 10 disk findings usually associated with glaucoma (tilted disk, saucerization, laminar dot sign, β zone atrophy, nasal cupping, barring of circumlinear vessel, notch, RNFL defect, and optic disk size). They also determined the horizontal and vertical cup-to-disk (C:D) ratio and made a final classification as glaucomatous optic neuropathy (GON) or normal. Their judgment was compared with that of a glaucoma expert (I.M.T.).

Interrater agreement for categorical variables was assessed by calculating the observed proportion of agreement (absolute agreement) and weighted κ statistics, for which 0.00–0.20 was considered poor, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 good, and 0.81–1 very good.⁵ For numerical variables (C:D), the mean difference and 95% limits of agreement (LOA) were calculated. A commercially available software was used for statistical analyses (Stata 12; StataCorp, College Station, TX, USA).

Results

Agreement between both groups of residents and the expert ranged from poor to moderate (Table 1).

Table 1 Agreement between ophthalmology residents and a glaucoma expert for ONH and retinal findings

ONH and retinal findings	First-Year Group	Third-Year Group
	Agreement	Agreement
	AA (%), κ	AA (%), κ
Tilted disk	87, $\kappa=0.24$	94, $\kappa=0.57$
Saucerization	67, $\kappa=0.05$	65, $\kappa=0.00$
Laminar dot sign	73, $\kappa=0.48$	80, $\kappa=0.60$
Zone β atrophy	55, $\kappa=0.16$	61, $\kappa=0.24$
Nasal cupping	60, $\kappa=0.23$	69, $\kappa=0.39$
Barring of circumlinear vessel	53, $\kappa=0.12$	63, $\kappa=0.28$
Notch	55, $\kappa=0.15$	64, $\kappa=0.15$
RNFL defect	52, $\kappa=0.03$	62, $\kappa=0.12$
Optic disk size	66, $\kappa=0.17$	62, $\kappa=0.19$
Glaucomatous neuropathy	64, $\kappa=0.19$	63, $\kappa=0.20$

Note: AA (absolute agreement): observed proportion of agreement.

Abbreviations: ONH, optic nerve head; RNFL, retinal nerve fiber layer.

All residents completed the evaluation proposed. Agreement between the Third-Year Group and the expert was better than that between the First-Year Group and the expert, especially in “nasal cupping”, “barring circumlinear vessel,” “tilted disk,” and “RNFL defect” criteria. The higher absolute agreement with the expert was found for the criteria tilted disk (87% in the First-Year Group and 94% in the Third-Year Group) and laminar dot sign (73% in the First-Year Group and 80% in the Third-Year Group). Excluding these characteristics, agreement with the expert in 8 of 10 criteria was between 52% and 70% in either group. Agreement with the expert in the detection of GON criteria was 64% ($\kappa=0.19$) for the First-Year Group and 63% ($\kappa=0.20$) for the Third-Year Group. The mean difference (95% LOA) between the First-Year Group and the expert in the vertical C:D ratio was -0.04 (-0.41 to 0.32) and that between the Third-Year Group and the expert was -0.04 (-0.35 to 0.26). The mean difference compared with the expert for the horizontal C:D ratio was -0.06 (-0.48 to 0.35) for the First-Year Group and -0.05 (-0.39 to 0.29) for the Third-Year Group (Table 2).

Discussion

The evaluation of ONH criteria is an essential skill in ophthalmologic practice. This is an enlightening study of changes in the ability to evaluate ONH criteria using stereo photographs and the C:D ratio during ophthalmology residency training. With regard to criteria for the diagnosis of GON, we found a weak to moderate agreement of both First- and Third-Year Groups with a glaucoma expert. Better agreement between the Third-Year Group and the expert was found for “nasal cupping”, “barring circumlinear vessel,” “notching”, and

Table 2 Agreement between ophthalmology residents and a glaucoma expert for C:D

C:D	Measure
Vertical C:D	
First-Year Group	
Mean (SD)	0.63 (0.21)
Mean difference (95% LOA) ^a	−0.04 (−0.41 to 0.32)
Third-Year Group	
Mean (SD)	0.62 (0.24)
Mean difference (95% LOA) ^a	−0.04 (−0.35 to 0.26)
Horizontal C:D	
First-Year Group	
Mean (SD)	0.63 (0.21)
Mean difference (95% LOA) ^a	−0.06 (−0.48 to 0.35)
Third-Year Group	
Mean (SD)	0.63 (0.23)
Mean difference (95% LOA) ^a	−0.05 (−0.39 to 0.29)

Note: ^aAgreement with the expert.

Abbreviations: C:D, cup-to-disk ratio; LOA, limits of agreement; SD, standard deviation.

“RNFL defect” criteria, which are the main ONH signs of glaucoma. We found 52%–70% agreement between both groups and the expert in 8 of 10 criteria, excluding tilted disk and laminar dot sign. With regard to the C:D ratio and the diagnosis of GON, we found similar agreement between either the two groups and the expert. Although the average difference in determining the relationship C:D ratio was almost nil, LOA varied widely, and this variation might be clinically relevant.

Previous studies that examined the effect of disk assessment on diagnostic accuracy have usually focused on assessment of a single disk characteristic, such as the vertical C:D ratio and whether the disk is glaucomatous or not. Azuara-Blanco et al⁶ examined the evaluation of whether the optic disk showed glaucomatous damage or was unchanged, although no specific instructions were provided about how to assess the glaucomatous changes of ONH. They found a fair to good interobserver agreement among glaucoma specialists, with κ values ranging from 0.34 to 0.68. Varma et al⁷ asked experts to determine the vertical C:D ratio and whether the disk was glaucomatous based on their clinical experience. No definition of glaucoma was provided. They found moderate interobserver agreement among experts in the estimation of the vertical C:D ratio (weighted κ [Kw] 0.67). Moreover, individual experts differed by as much as 0.16 disk diameters (DD). Abrams et al⁴ asked three groups to determine the C:D ratio and the presence or absence of glaucomatous damage using their own experience, and the participants were also not required to specify the criteria they used to arrive at their conclusion. This study showed significantly higher interobserver agreement for ophthalmologists (Kw 0.68) than

for optometrists (Kw 0.56) and residents (Kw 0.56) when estimating the C:D ratio. However, the interobserver agreement was only fair to moderate and was significantly lower than intraobserver agreement. This suggests that observers use different anatomic clues when determining the C:D ratio. Another study reported that ophthalmologists found disks of smaller size and those with milder glaucomatous loss the most difficult to classify correctly.⁸

However, the influence of other morphologic features has not been reported. A recently published article studied nine topographic features of the disk and RNFL in monoscopic optic disk photographs and found that moderate or large peripapillary atrophy, ovoid horizontal disk shape, and error in the assessment of the vertical C:D ratio, RNFL, cup shape (rim loss), or disk hemorrhage lead to underestimation.⁹ Our results agreed with the findings of this published study in terms of absolute percentages, but had lower κ coefficients. This was probably due to sample heterogeneity (κ paradox)¹⁰ and the specific glaucomatous ONH changes investigated in our study, which allowed for a better analysis of interobserver variability. Contrary to previous studies that focused on the analysis of one optic disk head characteristic at a time or used monoscopic photographs, our study conducted a broad analysis of 10 characteristics of the optical disk, besides the analysis of the C:D.

Potential limitations of our study include the limited variety of disk. Twenty disks may not fully reflect the phenotypic variation seen in a glaucomatous disk. The finding of disk hemorrhage was excluded from our analysis because no disk hemorrhage was present in the stereo photographs and our study was not designed to evaluate the ability of residents to identify disk hemorrhage. We also did not evaluate other signs, such as acquired pit and vessel diameter changes. Another limitation is that the findings by ophthalmology residents were compared with those of only one glaucoma expert. However, as has been documented by Varma et al,⁷ even expert evaluation of the C:D ratio can vary by as much as 0.2 DD monoscopically and 0.16 DD stereoscopically. Moreover, this was a cross-sectional study that compared two different groups of residents (first year versus third year) simultaneously, and, consequently, did not evaluate the same group longitudinally, which would be an adequate way of evaluating performance improvement. Also, the analysis of different groups of people can introduce a selection bias, for example, the first-year students are more efficient and therefore showed above average performance.

Performance evaluation of medical residents is important to analyze their improvement during training. Clearly, there is a need to optimize ophthalmologic teaching in such

important skills as identifying glaucomatous neuropathy. The results of this study may facilitate the creation of targeted teaching tools focused on parameters highlighted in the ONH.

Conclusion

ONH findings by the First-Year and Third-Year Groups showed weak to moderate agreement with those of a glaucoma expert. When comparing the two residents group, the Third-Year Group performed better only in a few criteria, primarily in the detection of “nasal cupping”, “barring circumlinear vessel,” “notching”, and “RNFL defect” criteria. This may suggest that the residents learn how to identify glaucoma signals during the first year of training.

Author contributions

All authors made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; took part in drafting the article or revising it critically for important intellectual content; gave final approval of the version to be published; and agree to be accountable for all aspects of the work.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Resnikoff S, Pascolini D, Etya'ale D, et al. Global data on visual impairment in the year 2002. *Bull World Health Organ*. 2004;82(11):844–851.
2. Caprioli J, Prum B, Zeyen T. Comparison of methods to evaluate the optic nerve head and nerve fiber layer for glaucomatous change. *Am J Ophthalmol*. 1996;121(6):659–667.
3. Zeyen T, Miglior S, Pfeiffer N, Cunha-Vaz J, Adamsons I; European Glaucoma Prevention Study G. Reproducibility of evaluation of optic disc change for glaucoma with stereo optic disc photographs. *Ophthalmology*. 2003;110(2):340–344.
4. Abrams LS, Scott IU, Spaeth GL, Quigley HA, Varma R. Agreement among optometrists, ophthalmologists, and residents in evaluating the optic disc for glaucoma. *Ophthalmology*. 1994;101(10):1662–1667.
5. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174.
6. Azuara-Blanco A, Katz LJ, Spaeth GL, Vernon SA, Spencer F, Lanzl IM. Clinical agreement among glaucoma experts in the detection of glaucomatous changes of the optic disk using simultaneous stereoscopic photographs. *Am J Ophthalmol*. 2003;136(5):949–950.
7. Varma R, Steinmann WC, Scott IU. Expert agreement in evaluating the optic disc for glaucoma. *Ophthalmology*. 1992;99(2):215–221.
8. Reus NJ, Lemij HG; European Optic Disc Assessment Trial (EODAT). Characteristics of misclassified discs in the European Optic Disc Assessment Trial (EODAT). *Invest Ophthalmol Vis Sci*. 2008;49(13):3627.
9. O'Neill EC, Gurria LU, Pandav SS, et al. Glaucomatous optic neuropathy evaluation project: factors associated with underestimation of glaucoma likelihood. *JAMA Ophthalmol*. 2014;132(5):560–566.
10. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol*. 1990;43(6):543–549.

Clinical Ophthalmology

Publish your work in this journal

Clinical Ophthalmology is an international, peer-reviewed journal covering all subspecialties within ophthalmology. Key topics include: Optometry; Visual science; Pharmacology and drug therapy in eye diseases; Basic Sciences; Primary and Secondary eye care; Patient Safety and Quality of Care Improvements. This journal is indexed on

Submit your manuscript here: <http://www.dovepress.com/clinical-ophthalmology-journal>

Dovepress

PubMed Central and CAS, and is the official journal of The Society of Clinical Ophthalmology (SCO). The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.